ED 155 891 ASPECTS OF LEXICAL SIMPLIFICATION IN THE SPEECH & WRITING OF ADVANCED ADULT LEARNERS. THE NOTIONS OF SIMPLIFICATION, INTERLANGUAGES & PIDGINS & THEIR RELATION TO SECOND LANGUAGE PEDAGOGY. ACTES DU 5EME COLLOQUE DE... LEVENSTON, E. A.; BLUM, S. 1977 24P.

1 OF 1

DOCUMENT RESUME

ED 155 891 FL 009 423

AUTHOR Levenston, E. A.; Blum, S.
TITLE Aspects of Lexical Simplification in the Speech and Writing of Advanced Adult Learners. The Notions of Simplification, Interlanguages and Pidgins and Their Relation to Second Language Pedagogy. Actes du 5eme Colloque de Linguistique Appliquee de Neuchatel (Acts of the Fifth Neuchatel Colloquium on Applied Linguistics).
INSTITUTION Neuchatel Univ. (Switzerland).
PUB DATE 77
NOTE 24p.; Paper presented at the Neuchatel Colloquium of Applied Linguistics (fifth, Neuchatel, Switzerland, May 20-22, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Communicative Competence (Languages); *Error Analysis (Language); Hebrew; Language Instruction; Language Research; Language Universals; *Learning Processes; Learning Theories; Linguistic Theory; *Second Language Learning; Semantics; *Vocabulary; *Vocabulary Development
IDENTIFIERS *Interlanguage; *Lexical Simplification (Language Learning)

ABSTRACT

This paper discusses the meaning of the term "lexical simplification" in the context of second language acquisition. It is suggested that simplification be viewed as a universal feature of language use which may be manifested in a number of linguistic contexts, including the creation of a learner's interlanguage. It is further suggested that studies of lexical simplification should concentrate on the process rather than its end product. A small-scale empirical investigation was conducted to examine a few selected aspects of the process of lexical simplification, especially those strategies which play a role in the formation of a learner's interlanguage and its subsequent fossilization. Eighty native speakers of Hebrew and 83 learners of Hebrew, on the university level, were given a set of 20 sentences containing missing words which the subjects were asked to provide. The learners were then asked to translate 30 Hebrew words and phrases into their mother tongue. The first instrument sought to measure lexical control in learners as compared to native speakers, while the second was intended to measure learners' ability to recognize lexical items as opposed to their ability to produce them. Results support the hypothesis of lexical simplification in language learners, as well as that of interlanguage fossilization. Test sentences used in the first part of the experiment are appended, in Hebrew with English translation. (AM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED155891

S. P. CORDER ET E. ROULET (ÉDS)

THE NOTIONS OF SIMPLIFICATION, INTERLANGUAGES AND PIDGINS AND THEIR RELATION TO SECOND LANGUAGE PEDAGOGY

ACTES
DU 5ÈME COLLOQUE
DE LINGUISTIQUE
APPLIQUÉE DE NEUCHÂTEL
20-22 MAI 1976

EXTRAIT – OFFPRINT

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Francoise Redard

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

FL009423

FACULTÉ DES LETTRES
NEUCHÂTEL

1977

LIBRAIRIE DROZ
GENÈVE

Aspects of Lexical Simplification in the Speech and Writing of Advanced Adult Learners*

E. A. LEVENSTON and S. BLUM, The Hebrew University of Jerusalem

1. Introduction

The simplest way of looking at lexical simplification is numerical: it is the process and/or result of making do with *less* words. And both the process and the result can be observed and studied in a number of different linguistic contexts:

i. in the speech and writing of second language learners – at all stages their total vocabulary is less than the vocabulary available to native speakers,
ii. in the speech (and, later, writing) of children acquiring their first language, before they possess the full range of vocabulary available to adult speakers,
iii. in the speech of adult native speakers when addressing children, foreigners (Ferguson 1971), learners of their language (Henzl 1973) or other native speakers they feel it necessary to "talk down" to because they possess a poorer vocabulary,
iv. in simplified reading texts prepared by native speakers for the use of second language learners (Blum 1972); the reduced range of vocabulary in which such texts are produced is often clearly specified by the publisher,
v. in pidgins (Samarin 1971: "the total lexical inventories of pidgins are known to be impoverished by comparison with those of natural languages"),
vi. in translation, when the lack of precise equivalents in the target language may lead to various devices similar to the kinds of lexical simplification found in the above situations (Dagut 1971).

Since lexical simplification is a feature of so many diverse linguistic activities, it seems reasonable to assume that it operates according to universal principles. And these universals of lexical simplification derive from certain aspects of *semantic competence*, particularly:

i. the awareness of hyponymy, synonymy, antonymy, converseness and possibly other systemic relationships between lexical items, by means of which we can explain why in specific contexts one lexical item can substitute for another (see also Richards 1976 and Slobin 1974),
ii. the ability to avoid the use of specific lexical items by means of circumlocution and paraphrase, and
iii. the ability to recognize degrees of paraphrastic equivalence.

* Both Larry Selinker and Andrew Cohen helped us immensely in the preparation of this report; as amateurs in psycholinguistic research, we appreciated their professional advice.

These aspects of semantic competence are the basis of communicative strategies that we use, when the need arises, to express complex meanings by indirect means with minimal vocabulary. For the language learner it is the difficulty, or impossibility, of both internalizing the exact nature of the inter-relationships mentioned above and acquiring the native speaker's awareness of degrees of polysemy and figurative extension and his sensitivity to formal, collocational and idiomatic restrictions on lexical choice that leads to continual use of simplification as a communicative strategy.

The learner's need to simplify is thus explained by the complexity of the task of acquiring command of all aspects of the native speaker's semantic competence; his ability to do so derives from his own semantic competence in his first language. As a native speaker he will at times experience a need for paraphrase and circumlocution, for hyponymy and synonymy, in some of the situations listed in the introductory paragraph above, (especially iii.). As a language learner he is compelled to reorganize semantic fields in the early stages, when he lacks vocabulary, and to do so according to the principles that govern his semantic competence as a native speaker.

At a later stage he may show signs of *lexical fossilization,* reverting to his first simplifications, or he may progress towards a semantic competence approximating that of a native speaker of the target language.

How can these phenomena of language simplification in second-language learners be studied? In principle there seem to be two alternative approaches, parallel to the ambiguity of the word itself:

1. "simplification" is understood as the *result* of simplifying, the end-product, the pattern of relationships displayed by the learner's total use of the language – see section 2.
2. "simplification" is understood as the *act* of simplifying, the strategy of communication, the process whereby specific meanings are communicated on specific occasions – see section 3.

2. Lexical simplification as end-product

An over-simple view of lexical simplification would concentrate on the *result* of making do with less, and view the task of research as lexicographic. For every word in the target language one would just have to list all the alternatives, single words or paraphrases, found in the learner's utterances. The end-product would be an "interlanguage dictionary". It would lack information, however, about the *processes,* or *strategies* of derivation whereby the alternatives were formed, or the structural relations between items in the list.

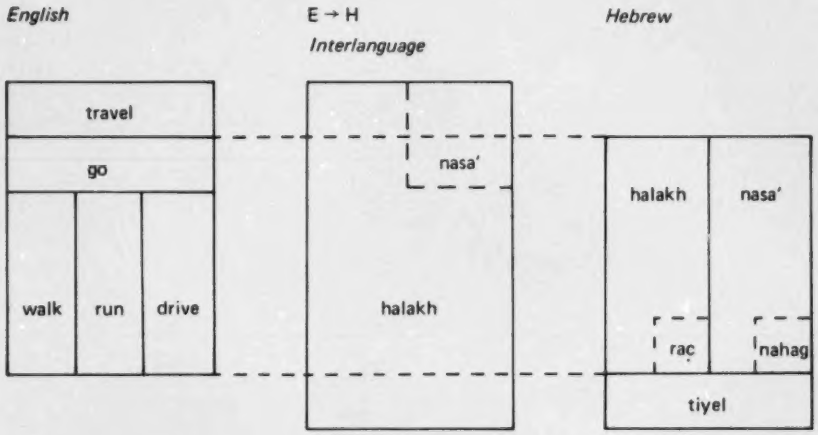
A subtler view of lexical simplification, though still concerned with an end-product, would attempt semantic "mapping", acknowledging that vocabulary cannot be adequately studied merely in terms of one-for-one equivalences. Instead of narrowing the focus to the study of lexical *items*, it would

compare broader *areas* of lexical teritory or – to change the metaphor and the number of dimensions – greater volumes of semantic space. For any given area – or volume – of vocabulary, it would be necessary to chart the way it was divided up a) in the learner's native language, b) in the learner's interlanguage and c) in the target language (cf. Selinker 1972: "the psychologically relevant data on second language learning are utterances in TL by native speakers, and in NL and IL by second language learners").

The main focus of attention would be the comparison of b) and c); comparison with the mother tongue should be explanatory, not predictive[1]. Such a study would be qualitative as well as quantitative. It would not merely demonstrate how lexical simplification meant managing with less words; it would also show how the range of uses of each word was affected by its place in the total structure of vocabulary.

A sample extract from such a study might look like this, using the technique of semantic mapping in two dimensions (For a fuller explanation of the details in the following diagram, see Levenston 1970: 222–4):

Figure 1.



1 The study of lexical errors made by Burmese speakers learning English (Myint Su, 1971 – as described in Pit Corder, 1973a) compared a) and b) but ignored the semantic mapping in English. Her results, therefore, demonstrate interference but do not speak directly to the question of simplification.

Alternatively, the same information for any word can be listed as a choice from a bundle of semantic components (marked as ± or unmarked when the distinction is neutralized):

	locomotive	*legs*	*engine*	*distance*
go	+			
halakh (Heb)	+	+	–	
nasa' (Heb)	+	–	+	
halakh (E → Heb. int.)	+			–
nasa' (E → Heb. int.)	+			+

Such techniques of representation have one great advantage – but several disadvantages. They do demonstrate graphically, for example, the way halakh (in the learner's speech) may be generalized to take over all or part of the field of nasa and tiyel (in the target language). They are strong on the study of referential meaning. But they fail to show any of the implications for lexical simplification of other aspects of vocabulary use. Words also differ in connotations, in collocability, in degree of formality, in other register restrictions, and in their ability to combine in idioms. And all these aspects of word usage are reflected in the learner's language. A word may be given not only a greater range of referential meaning; it may be used with inappropriate connotations, deviant collocations, too (in)formally, in the wrong register, unidiomatically.

Clearly an adequate description of lexical simplification must take account of all these dimensions of meaning. The mapping should take place in multi-dimensional space! And to the extent that information along any dimension is lacking for the first and second language – dictionaries are notoriously inadequate, for example, in their treatment of collocability – it will be necessary to include in the research program a study of native speakers' usage, observed or elicited, in both languages.

Even the data required for referential distinctions in the mapping is hard to come by. Much of it may be lying there in the dictionaries and thesauruses, but in completely unorganised fashion. And the lack of organization is a major obstacle to research.

Most of the work in the description of the structure of vocabulary remains to be done. Only the minutest areas have so far been staked out, and only one – the vocabulary of cooking – has been compared for English and Hebrew (Lehrer 1969, Newman 1975). As for the data required for mapping the interlanguage, it has first to be collected – by observation, elicitation or introspection. The examples given above were based on introspection i.e. largely guesswork.

3. Lexical simplification as process

So far we have discussed two views of lexical simplification as an end-product. Both of them are attempts to describe the result rather than the process. The first is inadequate and the second is impossible of achievement. Furthermore, they are both based on a rather shaky hypothesis: that the learner's interlanguage remains stable long enough for a description in detail to be feasible. As the major aim of research we would propose instead the investigation of the *process* of lexical simplification, especially the communicative strategies used by the second language learner.

Consider, for example, the way a learner copes with hyponymy in a second language. Semantic competence for a native speaker includes the recognition of relations of hyponymy between items in the lexical system. This competence enables the speaker a) to distinguish when a number of items have a common "superordinate" (hyperonym) which can replace them (flower ← rose) and b) to know the collocational/stylistic/semantic restraints upon such replacement.

In the second language, the lexical system may well be differently organized. There may be *no* superordinate term (wash//raḥaç, šataf , . .), or the distribution of items may require the use of a term not existing in the mother tongue (cut//ḥatakh, gazar). What happens when the learner, attempting communication, tries to organize for himself the system of hyponymy in the target language? He simplifies the system, in one or other of two major ways:

a) A "general" term – usually a common word, with a wide range of meanings, learnt in the early stages – is set up as a hyperonym, and used in contexts which require a more specific term. English-speaking learners of Hebrew attribute to sameaḥ contexts where "happy" or "glad" would be appropriate in English, but a more specific term is necessary in Hebrew: ani sameaḥ lagur be Yisrael (I am joyful to live in Israel).*

b) A specific term with clear semantic (and sometimes also morphological) restraints is raised in status and used as a hyperonym in all contexts e.g. lehagid is used for "speak", al tagid štuyot, geveret Levi higida bamištara (don't speak rubbish, Mrs. Levi told in the police station).

Although, to some extent, such utterances can be explained in terms of transfer from the mother tongue, it is also necessary to see the process as one of reorganising the lexical system of the target language.

Awareness of hyponymy is of course only one of the aspects of semantic competence that condition the learner's strategy of communication; converseness, antonymy and the various possibilities of circumlocution may also affect the way he utilizes a restricted vocabulary.

* Both here and later in section 4, the translations are attempts to achieve similar effects in English, and not accurate equivalents of the semantic *contents* of the original.

Two stages, we think, can be distinguished in the development of these processes of lexical simplification. In the first stage the learner lacks vocabulary, either active or passive, and is compelled to exploit to the maximum the limited lexical stock at his command. It is at this stage, for example, that – exploiting the principle of hyperonymy – he begins to use general terms in contexts where the native speaker would use a more specific word. In the second stage, the learner's stock of passive vocabulary has increased considerably. But his active use of the target language remains restricted; even where he is capable of distinguishing lexical uses and meanings in the speech of others, he fails to exploit these distinctions in his own speech.

Investigation of a learner's active vocabulary usage should always be accompanied, therefore, by a study of his passive knowledge. We need to know whether choice of a particular word is conditioned by complete ignorance or by passive knowledge of lexical alternatives. We need to know whether choice of a particular word is conditioned by complete ignorance or by passive knowledge of lexical alternatives. We also need to study why, in such circumstances, he prefers some words and ignores others. A number of hypotheses are worth investigation:

3.1. sequence and intensity of acquisition

The word learnt first and most intensively drilled remains dominant. This seems frequently to occur when the area of reference of a source language word is covered by *two* target language items e.g. English "live" = gar (demeurer), ḥayah (vivre) or English "time" = zman (temps), pa'am (fois). Certainly English speakers in their Hebrew do over-generalize the use of gar, which is also frequently practised paradigmatically as the standard example of its conjugation.

There are implications here for vocabulary selection in syllabus construction. The first candidates for over-generalization by the learner are often words learnt in the first stages of instruction; these were chosen for inclusion in the basic vocabulary *because of* their generality and ability to fit into many contexts. At the more advanced level, with the growing need for more precise expression and more powerful tools for the attainment of accuracy, this feature of generality – so helpful at the beginning – can become a positive hindrance.

3.2. *phonological avoidance.* Given the choice of two words or phrases sufficiently synonymous to be equally acceptable in many contexts, the learner will prefer the one he finds easier to pronounce. The word zi'za' (shock appal) will be replaced by hifḥid (frighten); hikhḥiš (deny) will be replaced by amar šelo . . / say that . . . not . . /.

3.3. *morphological avoidance.* The learner will prefer the regular verb to the irregular verb, whose forms he is not sure of. Of the two verbs which can translate "express" – bite' and hibi'a – the second is easier to conjugate in the past, perhaps because of its similarity to the common verb higi'a, and thus preferred. (Similarly, one of us suspects that in his French, he prefers "penser" in the past, though quite happy to use "croire" in the present).

3.4. *void avoidance.** Learners tend to avoid words for which no precise equivalents occur in their mother tongues, especially when the semantic components of such words require them to make distinctions they are not used to making at the level of single words. An example is the verb šibeç = to insert in a *suitable* place. This is replaced by hikhnis (insert) or sim (put) or by paraphrase.

3.5. *individual experience.* Room must be left for the possibility that individual learners develop their own idiosyncratic simplifications which will not fit any of the above hypotheses. This may be especially likely when the language is not acquired in a controlled classroom situation but by free contact with native speakers. In that case the kinds of context of situation in which the first words are acquired may influence the preferences which could later become "fossilized". In particular, the lack of exposure to certain registers or levels of formality leads to ignorance of register distinctions the native speaker takes for granted.

4. Communicative effects of simplification

It is also necessary, in the study of lexical simplification to examine the variety of effects it may have, as seen from the point of view of a native speaker trying to understand the message. Some uses can no doubt be classified as "errors", clearly inappropriate uses of one word in contexts where another is obviously required; anaḥnu yekholim ledaber 'al haba' ayot *šemaṣati* 'otan i.e. we can talk about the problems I came upon[1]. And some of these errors might be explained a posteriori by lexical contrastive analysis.

But such a combination of error analysis and lexical contrastive analysis a posteriori is doubly inadequate for the study of lexical simplification. First, in lexis as in grammar (Schachter 1974), it will not reveal *avoidance,* of whatever cause. Secondly, the term "error" is too broad, and fails to do justice to the wide range of phenomena produced by non-native choice of words. In addition to "recognizable errors", the following are some of the categories that seem to be needed in a subtler classification of these effects:

4.1. *inappropriate collocations.* The word chosen may be semantically, conceptually, appropriate but used in collocations that do not occur in native speakers' usage. The resulting utterance makes sense, but sounds strange – not so much an error, as the funny way foreigners use the language.

Example: letaqen ba 'ayot – to "mend" problems.

* A similar strategy is described under the name of "semantic avoidance" in Selinker, Swain and Dumas 1975.

4.2. *misunderstanding.* In other words, unrecognizable – or at least unrecognized – errors. For the linguist these are simply inappropriate lexical choices. For the listener they are words understood in a way different from that intended by the speaker. The misunderstanding may pass unnoticed, or it may lead to anything from puzzlement to armed conflict.

Example: ani sameaḥ lagur beYisrael – I am happy to stay (= live) in Israel.

4.3. *obscurity.* This seems particularly likely when a more general word, or hyperonym, is chosen instead of the specific, restricted term that native speakers would naturally prefer. The result is an acceptable sentence where reference is clearer to the speaker than to the listener.

Example: pa'am hayiti ro' eh 'oto harbeh, 'aval 'ani lo ro' eh 'oto bazman hazeh = I used to see him, but I haven't seen him this last time.

4.4. *ambiguity.* For the listener there is a difference between "obscurity", where he cannot fathom the speaker's intention at all, and "ambiguity", where he is not sure which of two or more clearly distinguishable meanings is the one intended. Ambiguity is likely when the learner over-generalizes the use of a common word without realizing its degree of polysemy. Like obscurity, it may be a feature of utterances that the native speaker, with a vaster range of more specific items, would be unlikely to produce.

Example: hi 'azva et habayit kedei leqabel samim = she left home to get drugs.

4.5. *inappropriate register or variety.* The word favoured by the learner as a general, all-purpose vocabulary item may be register-restricted in native speakers' usage. There are well-known examples of this phenomenon in pidgins and creoles e.g. the use of *ars* and *bagarap* in neo-Melanesian for "cause" and "be spoiled" (Hall 1966: 91–2). Both for pidgins and learners' interlanguage, the source of such choice is probably connected with the context of acquisition. If the first words acquired, for instance, are either formal or informal, it may be difficult for the learner later to acquire those lexical distinctions which are register-based and not conceptual.

Example: ani ḥafeç kos teh = I desire a cup of tea

In principle lexical items from any variety of the language – literary, dialectal, technical, sub-standard – could be over-generalized. In practice it seems to be the level of formality which gives rise to most incongruities. A frequent feature of learners' speech is what might be termed "register-clash', where words are used from different registers and it may not be clear from the context which was the register the speaker was aiming at, or even whether he was aware of their existence.

Example: ani 'omed 'al kakh lir' ot 'oto, im zeh beseder 'itkha = I gotta see him, with your gracious permission.

5. Experimental Report

Of the various aspects of lexical simplification suggested in the preceding discussion we chose four for empirical investigation. Formulated as research hypotheses about the process of simplification in the learner's speech or writing, they read as follows:

a. The learner will *generalize* the use of one word in contexts where the native speaker will use different items of more restricted applicability (see the discussion of hyperonymy above p. 55).
b. The learner will use words of wider use, unmarked for *register,* in contexts where the native speaker will feel the need for register-restricted lexical items.
c. The learner will overgeneralize relationships of *antonymy,* using both members of a pair of antonyms in contexts where the native speaker will use specific, context-restricted lexical items.
d. Even when the learner does know a fixed *collocation* he will tend to use a word of more general applicability where the native speaker will more frequently choose the fixed collocation.

A necessary feature of the investigation, already clear from the way the hypotheses have been formulated, was the direct comparison of native speakers' usage with that of learners.

The lexical items to be studied were chosen on the basis of experience teaching Hebrew at the advanced level; we sought empirical confirmation of teachers' hunches concerning both learners' errors and native usage.

Previous investigations of similar problems had used translation into the target language (Taylor 1975) or pictorial stimuli (Selinker et al., 1975). Translation was inappropriate to our needs, since it would not allow adequate comparison with native speakers' usage. We decided against pictorial stimuli as likely to prove too open-ended. An elicitation procedure thought to provide both comparative data from native speakers and relatively focussed responses was that of sentence completion, whereby a single lexical item – word or phrase – would have to be inserted in each sentence. The study thus also tests the validity of the sentence completion procedure for this type of research.

5.1. Instruments

Two instruments were devised for the research:

a) A set of twenty sentences (see Appendix A), listed in random order, designed to test the four hypotheses i.e. generalization (sentences 1,5, 10, 14, 17), antonymy (sentences 3,7,8,11) register (sentences 2,4,6,9,18,19) and collocation (sentences 13,16,20). Two further sentences (12 and 15), of no experimental significance, were added to bring the number up to 20 and to help obscure the purpose of the experiment from the subjects. The instructions were: "Please fill in the missing words. Always write down the first word that comes to mind". This instrument was given to both learners and native speakers.

b) A list of thirty words and phrases which learners were asked to translate into their mother tongue, including the twenty hypothesized as most likely to be used by native speakers in the test sentences. This instrument was given to learners only; it was designed to compare the learner's ability to *recognize* lexical items with his ability to *produce* them. Since the sentence-completion exercise was administered immediately before this word-translation test, the ten extra words were included to reduce the learning effect of the first instrument upon the second.

5.2. Subjects

The subjects were 80 native speakers of Hebrew and 83 learners. All were university students, 20–30 years of age, male and female; the possibility of sex being a significant variable in this research was not investigated.

Of the native speakers the majority were undergraduates studying English as a foreign language in summer courses. The remainder were graduate students who were present at the annual conference of the Israel Association of Applied Linguistics.

Of the learners, 41 were native speakers of English and 42 were speakers of other European languages (16 Spanish, 8 Russian, 8 French, 5 Italian, 3 German, 2 Portuguese). The sample was structured in this way in order to compare the results of a group with a single mother tongue, English, with those of a group of mixed language backgrounds and thus see how far the influence of a specific mother tongue is significant in the processes of language simplification.

A preliminary study of the results, however, revealed no obvious differences and we decided to continue the research in three stages. In the first stage, reported here, we have compared the native speakers with *all* learners, analysed as a single group. In the second stage we shall compare English speakers with a mixed group of learners, as suggested above. In a third stage we shall obtain equal-sized groups for each mother tongue, and do a more thorough series of contrastive studies of mother-tongue influence on lexical acquisition and simplification.

As far as the level of proficiency in Hebrew was concerned, 50 of the learners were students in university summer courses in Hebrew, and had been assigned by placement tests to classes five and six i.e. the highest levels; by the end of the course they would be candidates for exemption from the need for any further formal study of Hebrew. The remaining 30 were at the end of one year's study of Hebrew at the university, also at the highest level. No attention has yet been paid to the length of time spent by the subjects in Israel, but very few were without considerable experience of living in Israel; they were all clearly "advanced" learners.

5.3. Procedures

The tests were administered to the subjects in groups, usually at the beginning or end of a period of class instruction (in Hebrew or English as a

foreign language). Either one of the principal investigators or their research assistant was present on all occasions to answer queries. Subjects were instructed not to complete any sentence which was not fully understood. Hebrew was used to administer the tests; all learners were sufficiently advanced to follow instruction, written and spoken, in Hebrew.

5.4. Results and Analysis

This section presents the results for 12 of the 18 test sentences i.e. all those where the majority of native speakers did choose the lexical item expected by the investigators, and where in most cases there was no single alternative used by a significant number of the remaining native speakers. Where there was a single alternative chosen by native speakers, notably in the sentences designed to test the *register* hypothesis, it was always the word which learners were expected to produce. The significance of these responses is discussed below for each hypothesis where they occurred. 6 sentences had to be rejected; they failed to produce sufficient agreement by native speakers as to the most likely completion. There seem to have been three reasons for this failure:

(1) The sentence was not marked sufficiently clearly as *register-restricted,* and the register-specific item expected was therefore not always chosen – see no. 19.
(2) The context was not defined with sufficient specificity, allowing for a wider range of meanings than expected – see no. 2.
(3) Even where the context was clearly defined, a greater range of *synonyms* proved possible than we expected. It should be pointed out that pilot testing had been carried out with most of the sentences, and the versions quoted here are often the revised, presumably improved, versions of earlier failures; even so, not all the revisions worked. Anyone contemplating further research along similar lines is warned of the difficulties associated with constructing test sentences that produce unanimous responses from native speakers. Quite apart from the freedom of imagination with which some subjects will always treat the most context-restricted sentence, there will also be others who demonstrate lexical simplification and choose more general words than those expected. This was an unexpected but rewarding feature of the experimental results; the proof that lexical simplification is used by native speakers.

The following procedures have been followed in displaying the data:
(1) The result for the 12 sentences, 3 for each hypothesis, are presented in 4 separate sets of tables,
(2) Taking a sample item from sentence 1 (see below), the data entries in the first row represent the percentage of subjects within a column who used the word expected from native speakers.

Sentence 1	Natives n=80	Learners n=83	
		+ n=53	− n=30
meušar, word expected from natives	80.0	28.3	3.3
sameaḥ, word expected from learners	2.5	33.4	43.3
other	17.5	37.7	53.3

(3) The data entries in the second row in Tables I, III and IV represent the percentage of subjects who used the word expected from learners, the response pattern which supports the research hypothesis. (For the second and third rows in Table V see p. 67).

(4) The data entries in the last row in Tables I, III and IV represent the percentage of subjects who made some "other" response. There is however a difference between "other" responses of native speakers and those of learners. "Other" responses by native speakers were all acceptable, and demonstrated their ability to find other solutions, with considerable change of meaning, than those expected by the investigators. "Other" responses by learners includes both acceptable and unacceptable sentences. More research is needed into the variety of unacceptable sentences produced.

(5) The tables offer a further division of learners into those who knew or did not know the expected word, as determined by their performance on the translation test. For each sentence, the subgroup who knew the word are represented by a "+" and those who did not by a "−". Thus in our example above 53 learners knew meušar in translation, whereas 30 did not.

(6) All data are expressed as percentages, even when the size of the group is too small for a percentage figure to have much significance e.g. sentence 7, where only 3 learners did not "know" the expected word, (i.e. translate it accurately in the translation test) though one used it.

Although the results will also be reported hypothesis by hypothesis, the following general points apply to all the data produced by learners:

(1) All learners, whatever their language background, had a tendency to ignore the grammatical restraints of concord in choosing lexical items, e.g. in sentence 6 they used a feminine noun, though the verb is marked as masculine.

(2) With one exception, there is always a considerable gap between a learner's knowledge of a word and his use of the word in the appropriate context. This gap becomes apparent in comparing the percentage of Israelis using the expected word with the percentage of learners who both knew it and used it. The single exception is sentence 3, where the learners "behave better" than the native speakers. The explanation here probably lies in the wider range of acceptable collocations for "money" known to native speakers.

(3) The results for all four hypotheses could be conflated into one single hypothesis which seems to hold good for the interlanguages of the learners studied; the learner will use words of more general meaning and applicability to a greater extent than the native speaker.

(4) Learners are less sensitive than native speakers to the contextual constraints of semantic choice imposed by the rest of the sentence. This is always clear from the last column of each table, where the percentage of "other" learner responses is nearly always higher, especially for those subjects who did not know the expected word.

5.4.1. Generalization

In the usage of native speakers of Hebrew, the vocabulary items designed to test the generalization hypothesis seem to be distinguished as follows:

sameaḥ – "joyful", "glad" i.e. "happy" in a way that is expressed *physically.*
meušar – "happy", "content" i.e. "happy" as descriptive of a spiritual state that may or may not be physically recognizable.
meruçe – "content", "satisfied" i.e. "happy" in the sense that all demands have been met.

Table I: Data for the sentences testing the generalization hypothesis

Sentence 1	Israelis n = 80	Learners n = 83	
		+ known n = 53	– unknown n = 30
word expected from native: meušar	80.	28.3	3.3
word expected from learners: sameaḥ . . .	2.5	33.4	43.3
other	17.5	37.7	53.3
Sentence 14	n = 80	n = 70	n = 13
word expected from natives: meruçe	91.2	42.8	–
word expected from learners: sameaḥ . . .	1.25	18.6	15.4
other	7.5	38.6	84.5
Sentence 17	n = 80	n = 80	n = 3
word expected from natives: sameaḥ . . .	75.	48.7	33.3
synonym	11.25	10.	33.3
other	13.75	41.2	33.3

Our hypothesis was that learners would tend to generalize the use of sameaḥ in all sentences. This hunch is clearly supported, but since sentence 17 required the use of sameaḥ from native speakers as well, more interesting results were obtained by a more in-depth analysis combining the figures for sentences 1 and 14. This is done in the following table, where knowledge of the Hebrew words is further analysed as follows:

column 1 -- native responses.
column 2 – response patterns of subjects who knew one word only, in all cases sameaḥ.
column 3 – response pattern of subjects who knew two words only of the three expected, generally sameaḥ and meruṣe, but in 9 cases sameaḥ and meušar.
column 4 – response pattern of subjects who knew all three words.

Table II: Distribution of responses in sentences 1 and 14

word	Israelis n = 160	Learners knew 1 word n = 22	knew 2 words n = 30	knew 3 words n = 114
sameaḥ	1.8	36.3	33.3	23.7
meušar	41.9	4.5	10.0	14.0
meruṣe	45.0	–	13.3	30.7
svea' raṣon	1.8	–	–	1.75
mabsut	1.2	–	6.7	3.5
other replies	6.9	4.5	23.3	14.0
no reply	–	54.5	13.3	11.4

The process of lexical fossilization is made beautifully clear by the first row of data. Though the percentage of subjects using sameaḥ steadily declines with increasing knowledge of vocabulary, the choice of sameaḥ remains prominent, even when all three items are known.

Learners seem to continue using words of more general meaning and applicability, in acceptable and unacceptable contexts, even when aware of other more specific alternatives. The difference between this kind of simplification by learners and the native speakers' use of the same strategy is both quantitative (frequency of occurrence) and qualitative (degree of acceptability). The semantic structure of the interlanguage is thus different from that of the target language: compare, for example, the diagrams suggested in section 2, p. 53 above).

* of the other two items specified in this table, mabsut is an Arabic loan word, common in informal speech, and svea' raṣon is a very close synonym for meruṣe – the reason for its avoidance by native speakers requires further investigation.

5.4.2 Register

In the usage of native speakers of Hebrew, the vocabulary items are distinguished as follows:

mavet = "death", unmarked for register.
ptira = "decease", used for "death" in newspaper announcements of bereavement, and other contexts requiring euphemism.
ḥika = "wait", unmarked for register.
ḥimtin = "wait", in telephone communication (hamten 'al haqav = "hold the line! " – lit. "wait on the line") and other formal contexts.
raḥaç = wash (any part of the body), unmarked for register.
natal = wash (the hands), as part of religious ritual.

Table III: Data for the sentences testing the "register" hypothesis

	Israelis n = 80	Learners n = 83	
		+ known	– unknown
Sentence 9	n = 80	n = 72	n = 11
word expected from natives: ptira	75.	12.5	–
word expected from learners: mavet ..	22.5	63.8	36.4
other	2.51	23.6	63.6
Sentence 4	n = 80	n = 55	n = 28
word expected from natives: natal	85.	38.2	3.6
word expected from learners: rahac....	12.5	60.	82.1
other	2.5	1.8	14.3
Sentence 18	n = 80	n = 43	n = 40
word expected from natives: himtin ...	85.	27.9	2.5
word expected from learners: hika	12.5	27.9	32.5
other	2.5	44.2	65.

The high percentage of learners in the second row of data for each sentence clearly supports the hypothesis: in sentence 9, for example, clearly marked as a newspaper announcement, 46 of the 72 learners who recognized the word ptira/decease still used the general word mavet. Particularly interesting here, however, is the fact that the same choice was made by 18 of the 80 native speakers. And the same phenomenon occurs, though less markedly in the other two sentences. It may be that not all native speakers of a language are equally sensitive to register distinctions within that language; to this extent their semantic competence is limited.

Another way of looking at this result is to regard it as further evidence of lexical simplification as a communicative strategy used by native speakers. At all events, one implication of the research is the need for studying register sensitivity among native speakers; we have found no examples in the literature of any such studies.*

5.4.3. Antonymy

The pair of antonyms examined was natan/qibel = give/receive (also accept); one member of the pair appeared in each of the sentences devised to elicit the other.

Table IV: Data for the sentences testing the "antonymy" hypothesis

	Israelis n = 80	+	learners n = 83	–
Sentence 7	n = 80		n = 80	n = 3
word expected from natives: masar	77.5		48.7	33.3
word expected from learners: natan	–		5.2	–
other more general words	13.7		28.7	33.3
other responses	8.3		16.2	33.3
Sentence 11	n = 80		n = 61	n = 22
word expected from natives: nicel	88.7		59.	13.6
word expected from learners: qibel	1.2		4.9	18.2
other more general words	–		16.9	63.6
other response	10.		18.3	4.5
Sentence 3	n = 80		n = 53	n = 30
word expected from natives: hirviah	67.5		79.2	13.3
word expected from learners: qibel	8.7		13.2	56.6
other more general words	10.		–	6.6
other responses	19.7		7.5	23.3

On the whole the results do not support the hypothesis for sentences 7 and 11; the percentage of learners using qibel/natan, especially of those who knew the specific term, is quite low. But some did; resorting to antonymy is clearly one possible strategy, though not as common as we had hypothesized. What did

* There have been studies where subjects have been asked to classify usage as "acceptable/unacceptable in formal/informal speech/writing" e.g. Mittins et al. *Attitudes to English Usage*, O.U.P., 1970, but none where subjects were required to actually produce the register-marked items.

emerge clearly from the data, however, was the continued tendency – as shown by the third row of data in Table 4 – to use a more general term even when the specific item was known e.g. in sentence 7 higid = say rather than masar = transmit – in sentence 11 hištameš = use or laqaḥ = take, rather than niçel = exploit.

In sentence 3 the word expected from learners, qibel, was used by a majority of the learners who did not recognize hirviaḥ.** However, qibel is a word of very wide range, and this result could just as well be interpreted as further evidence for generalization as a strategy in lexical communication.

Note that qibel also occurs more frequently here than in sentences 7 and 11 among native speakers; the collocation exists. One effect of this may be that it is much less likely to be corrected by a teacher when used by a learner, even when a more specific word is *more* acceptable, and hence becomes fixed as an all-purpose expression among language learners.

5.4.4. Collocation

In Table V, "other collocations" means for native speakers the use of other synonyms which are equally acceptable, and do not change the meaning: ne'evaq (struggle) in sentence 13 and mile' (fulfil) in sentence 20. There were no alternatives in sentence 16.

Table V: Data for the sentences testing the "collocation" hypothesis

	Israelis n = 88	Learners n = 83	
		+	–
Sentence 13	n = 80	n = 53	n = 30
word expected from natives: hitmoded. . . .	82.5	25.4	3.3
other collocations	2.5	39.6	46.6
other responses	15.	34.	50.
Sentence 16	n = 80	n = 73	n = 10
word expected from natives: mabat rišon	86.2	42.4	50.
other collocations	–	30.7	20.
other responses	13.7	27.4	60.
Sentence 20	n = 80	n = 53	n = 30
word expected from natives: qiyem	77.5	41.5	13.3
other collocations	17.5	15.5	10.
other responses	5.4	43.4	76.6

** The high percentage of subjects, 13%, who did use hirviah, even though they could not translate it, is explained by spelling difficulties; they all misspelt it.

For learners, "other collocations" means any combination of words which has the same meaning as the expected collocation, regardless of whether it would be considered by a native speaker as acceptable or unacceptable e.g. behistaklut rišona ("on first looking") in sentence 16, šamar havtaḥa ("preserve a promise") in sentence 20.

The high number of learners who were able to translate 'ahava bemabat rišon (love at first sight) is probably due to its transparency and/or the fact that it is a loan translation. They may not all have come across it before the test, and it probably should not have been included. In sentence 13 hitmoded (lit. measure oneself against) and in sentence 20 qiyem ("keep" a promise) were far less likely to be translated unless previously encountered.

For the sentences reported on in Table 5 there is no word expected from learners; what is of interest is the range of strategies used by learners who do not know or cannot recall the appropriate collocation. Most of the alternatives chosen – and represented by the entries in the second row – were more general in use than the expected words preferred by native speakers, and thus do support the hypothesis. However, it would take up too much space to demonstrate this in detail: a large number of words were used, some of them not more than once or twice.

6. In conclusion

In the first part of this paper – sections 1–4 – we have discussed the meaning of the term "lexical simplification" in the context of second language acquisition. We suggest that simplification be viewed as a universal feature of language use which may be manifested in a number of different linguistic contexts, including the creation of a learner's interlanguage.

The probable universal nature of lexical simplification is supported by two arguments:

i) the *need* is constant, deriving from limits placed on the size of the vocabulary at the disposal of the speaker. These limits may stem from the *speaker's* own lack of knowledge, the addressee's linguistic immaturity, or – in the case of pidgins – features of the language itself;
ii) the *strategies employed* by a speaker to achieve his communicative goals despite such restraints are based on certain aspects of semantic competence shared by all speakers.

In the specific case of second language acquisition, we have tried to distinguish between those aspects of semantic structure – hyperonymy, antonymy, converseness, synonymy, etc. – which make lexical simplification possible, and those complexities of language use – register-variety, metaphorical extension, collocational restriction, etc. – which make it *necessary*. However, though hyperonymy, antonymy, converseness and synonymy are

universals of semantic structure, their realization will differ from one language to another, and thus the *end-product* of lexical simplification in a learner's speech will not always be acceptable to a native speaker.

Investigation of lexical simplification should concentrate, we believe, on the study of the *process* rather than the description of the *end-product*. Among the factors that lead to the creation of a lexically simplified interlanguage and may contribute to its subsequent fossilization, we suggest the following as worthy of study in detail: the sequence and intensity of acquisition, phonological avoidance, morphological avoidance and "void" avoidance. The communicative effects upon native speakers of lexical simplification by learners should also be studied. In addition to the clear recognition by the listener of referential errors and inappropriate collocations or registers, there will also be cases of misunderstanding, obscurity and ambiguity.

The second part of this paper reports a small-scale empirical study designed to investigate a few selected aspects of the process of lexical simplification. We chose to concentrate on strategies which would help us understand both the formation of a learner's interlanguage and the reasons for its subsequent fossilization. In this connection, it is necessary to distinguish between the obligation to use lexical simplification as a communicative strategy when the learner has no choice but to make do with fewer words, and the tendency to continue resorting to these simplifications even when the need for them no longer exists.

From the data summarized in the research report, it seems clear to us that the research technique we have devised is a valid and useful method of investigating lexical acquisition and simplification. The findings probably have pedagogical implications both for vocabulary selection in materials production and for classroom practice, but this must await further results. A number of unresolved problems that arose during the course of the experiment also await further research; the relevance of lexical contrastive analysis, the nature of register sensitivity among native speakers. The data presented in the last section represent a promising beginning, but much remains to be done. Other hypotheses among those proposed in section 3 could be tested by similar methods e.g. phonological avoidance, "void" avoidance. Even the hypotheses described as supported require more evidence with other sets of words. And, of course, with other pairs of languages.

Appendix A – Test sentences

1. tov lahem beyaḥad, ani ḥošev šehem me' od ———— benisu' ehem.
2. hašofet ———— lehaḥzir šne yeladim le' imam ha' ameriqa' it.
3. hem natnu lo 'avoda tova, hu ———— harbe kesef.
4. lefi hamesoret hayehudit, ———— yadayim lifne ha' aruḥa.
5. kešešama šenolad lo ben hu haya me' od ———— .
6. maḥar yitqayem ———— yehudi-'aravi benaçrat.

7. ani miçta 'er šehu lo qibel 'et drišat hašalom mimxa, pašut šaxaḥti ———— .
8. hu haya muxan leqabel beraçon' et ha 'aḥarayut ———— 'alav.
9. anu modi"im beça'ar rav' 'al ———— šel dodatenu hayeqara Mazal Levi, halevaya titqayem maḥar . . .
10. hu roçe la'azov 'et ha'areç; qase lo limço 'avoda tova, ve'en lo ḥaverim. beqiçur, ani ḥošev šehu lo ———— kan.
11. ani ḥošev šenatnu lo maspiq hizdamnuyot. haçara hi šehu lo yode'a' ex ———— 'otam.
12. zarqu'oto meha'universita ki hu ———— bexol habḥinot.
13. 'asur livroaḥ miba 'ayot; çarix ———— 'itan.
14. hu qibel misra ḥadaša vuhu haya me' od ———— mimena.
15. hu lo limed bixlal, qaše li ———— še yaçliaḥ.
16. qaše li lena' amin be' ahava me ———— 'anašim çrixim lehakir ze et ze kede lehit' ahev.
17. hu'ohev anašim: hu tamid ———— kešeba' im levaqer 'oto.
18. hamerkazanit hitnaçla 'al kax šelo hayu qavim pnuyim vebiqša mimeni ———— zman ma.
19. harov 'azvu. hayom ———— bilvanon raq 500 yehudim.
20. lo kedai lexa lifnot 'elav. hu lo ———— 'et havtaḥotav.

Appendix B – Test sentences in translation

(N.B. The following translations provide idiomatic English equivalents of the Hebrew sentences, but would not necessarily serve as an equivalent research instrument for studying simplification in English).
Please fill in the missing words. Always put down the first word that comes to mind.

1. They get on well together. I think they have a very ———— marriage.
2. The judge ———— that the two children be returned to their American mother.
3. They gave him a good job. He ———— a lot of money.
4. According to Jewish tradition, one must ———— one's hands before a meal.
5. When he heard that he was the father of a son, he was very ————.
6. Tomorrow an Arab-Jewish ———— will take place in Nazareth.
7. I'm sorry but he did not receive your kind regards. I simply forgot to ————.
8. He was ready to accept the responsibility that had been ———— upon him.
9. We deeply mourn the ———— of our beloved aunt, Mazal Cohen. The funeral will ————.
10. He wants to leave Israel. He can't find a good job, and he has no friends. In short, I don't think he's ———— here.
11. I think he's been given enough opportunities. The trouble is that he doesn't know how to ———— them.
12. They expelled him from the university because he ———— in all the exams.
13. Problems must not be shirked; they must be ————.
14. He took a new job and was very ———— with it.
15. He hasn't learnt a thing. I find it hard to ———— that he'll succeed.
16. I find it hard to believe in love at ————. People have to know each other in order to fall in love.
17. He likes people; he is always ———— when they come to visit him.
18. The operator apologized for the fact that all the lines were engaged and asked me to ———— for a while.
19. Most have left. Today there are only 500 Jews ———— in Lebanon.
20. It's not worth turning to him. He doesn't ———— his promises.

References

Blum, Shoshana (1972): *Easy Hebrew as a means for teaching Hebrew to adults.* Unpublished Ph. D. Thesis, the Hebrew University of Jerusalem.

Dagut, Merton (1971): *A linguistic analysis of some semantic problems of Hebrew-English translation.* Unpublished Ph. D. thesis, the Hebrew University of Jerusalem.

Ferguson, Charles A. (1971): "Absence of copula and the notion of simplicity: a study of normal speech, baby talk, foreigner talk and pidgins", in Hymes (1971).

Hall, Robert A. Jr. (1966): *Pidgin and Creole Languages,* Ithaca, Cornell University Press.

Henzl, M. Vera (1973): "Linguistic Register of Foreign Language Instruction", *Language Learning* 23/2, 1973.

Hymes, Dell (ed.) (1971): *Pidginization and Creolization of Languages,* Cambridge, at the University Press.

Lehrer, A. (1969): "Semantic cuisine", *Journal of Linguistics* 5.

Levenston, E. A. (1970): *English for Israelis: A guide for teachers,* Jerusalem, Israel University Press.

Myint Su (1971): *The Analysis of Lexical Errors,* unpublished M. Litt. Thesis, University of Edinburgh.

Newman, Arieh (1975): "A semantic analysis of English and Hebrew Cooking Terms", *Lingua* 37.

Corder, Pit (1973a): "The Elicitation of Interlanguage", in *Errata, Papers in Error Analysis,* ed. Jan Svartvik, Lund.

– (1973b): *Introducing Applied Linguistics,* London, Penguin.

Richards, Jack C. (1976): "The Role of Vocabulary Teaching", *TESOL Quarterly* 10 (1).

Samarin, William J. (1971): "Salient and Substantive Pidginization", in Hymes (1971).

Schachter, Jacquelyn (1974): "An error in error analysis", *Language Learning* 24/2.

Selinker L., Swain M., and Dumas G. (1975): "The interlanguage hypothesis extended to children", *Language Learning* 25/1.

Slobin, Dan. I. (1974): *Psycholinguistics,* London, Scott, Foreman.

Taylor, Barry (1975): "The use of overgeneralization and transfer learning strategies by elementary and intermediate students of English as a second language", *Language Learning,* 25/1.

A Model for the Interpretation of Utterances

Margareta OLSSON, Lärarhögskolan, Mölndal

1. Background

Nowadays we emphasize communication as the foremost goal of foreign language instruction. In these circumstances the lack of guidelines for an assessment of erroneous utterances from a communicative point of view is a serious drawback. In order to find such guiding principles I will attempt below to establish what decides the communicative value of sentences while making use of earlier findings and general knowledge about linguistic matters. By communicative value I mean roughly the degree to which a speaker succeeds in conveying the intended information to a receiver.

2. Point of Departure

When confronted with an overtly erroneous utterance in an everyday situation a listener or reader will attempt to impose an interpretation on it. It is evident, too, in the extensive body of literature on the subject that testees try to reconstruct or interpret imperfectly delivered utterances according to instructions. Moreover, the documented results in these studies suggest that these activities do not proceed in a haphazard fashion. On the other hand, the results do not suggest a universally applicable explanation about what makes certain utterances easy to interpret whereas others resist interpretation to a very high degree.

3. On Grammatical and Lexical Errors

A receiver's reconstruction of a deviant utterance may consist of either a grammatical or a lexical change. A grammatical change would involve, for example, morpholigical inflections, derivations, function words, intonation, and word order. A lexical change would in general involve the word base of a content word.

In all decoding activities the general context (linguistic as well as situational) plays a primary role. Thus the choice of form as well as content is guided by considerations of overall acceptability. To illustrate, the lexically deviant utterance *The pheasants had started harvesting* and the syntactically deviant utterance *The lion was devoured by the carcass* will without much difficulty yield the intended interpretations *The peasants had started harvesting* and *The carcass was devoured by the lion.* This is because the function of language is, after all, communication, and we consequently expect the